数理解析研究所講究録
353巻 1979年 1-10

Representability Problem for Relational Database Design with Multivalued Dependencies

College of Liberal Arts, Kobe University
Katsumi TANAKA
Dept. of Information Science, Kyoto University
Yahiko KAMBAYASHI
Shuzo YAJIMA

1. Introduction

In the relational database model introduced by Codd[1], data values are collected as tables called relations. The columns of a table are labelled by distinct attributes. Several types of dependencies about attribute relationships have been introduced to specify the intensional properties of a relation, such as functional dependencies(FDs)[1] and multivalued dependencies(MVDs)[2].

One essential problem of the relational database design is the representability problem in the normalization[1,2], that is concerned with how much the designed relation schemata (table skeltons) represent the intensional properties of data.

Recently, Beeri et al. provided several criteria for the representability problem[3]. One of the criteria is as follows: A set of relation schemata represents the same information as a given initial relation schema if the former has the same dependencies and the same data as the latter. In [4], Rissanen showed a necessary and sufficient condition for a set of relation schemata to satisfy this criterion when only FDs are involved. When MVDs, a generalization of FDs, are also involved, Rissanen's result cannot be directly applied because of the difference between FDs and MVDs.

In this paper, we provide a new criterion for the representability of the decompositional schema design of a relational database when MVDs and FDs are involved as constraints. The representability problem including MVDs is concerned with lossless join[3,5], covering for MVDs and FDs, and preserveness of MVDs under update operations. In order to handle these problems, several properties of embedded multivalued dependencies[2] are first investigated. By the results, we introduce two independent properties for a necessary and sufficient condition of the representability, and provide the methods to examine them.

2. Basic Concepts

A relational database schema consists of a finite set of relation schemata, a set of domains and a set of integrity constraints. Each relation schema consists of a schema name, together with a finite set of attributes. A relation schema is denoted by $R(A_1,\ldots,A_n)$ or $R(U_i,V_i,W_i,X_i,Y_i,Z_i)$ etc., where R is a schema name, uppercase letters from the begining of the alphabet are names for attributes, and those from the end of the alphabet are names for sets of attributes. An instance of $R(A_1,\ldots,A_n)$ is a finite mathematical n-ary relation, which is a subset of the Cartesian product $\prod_{i=1}^{n} D(A_i)$. Here, D is a mapping from the set of attributes to the set of domains. When no confusion occurs, R is used to denote an instance of the schema and called a relation.

$R[X]$ denotes a projection[1] on a set X of attributes of R. $R*S$ denotes a natural join[1] of relations R and S. Following Fagin's notation[2], we denote $Z_R(x)=\{r[Z];\ r[X]=x,\ r\in R \text{ and } X,Z\subseteq U\}$.

A multivalued dependency (for short, MVD) $X\twoheadrightarrow Y$ holds for a relation schema $R(U)$ iff in every allowable instance R, $Z_R(x)=Z_R(xy)$ holds for any XY-value xy in $R[XY]$. Here, X,Y and Z are sets of attributes such that $X\cup Y\cup Z=U$. In this paper, the set union of X and Y is sometimes denoted by XY for short. When the MVD $X\twoheadrightarrow Y$ holds for $R(X,Y,Z)$, we also denote it by $X\twoheadrightarrow Y|Z$ since $X\twoheadrightarrow Y$ implies that another MVD $X\twoheadrightarrow Z$ complementarily holds for $R(X,Y,Z)$. If Y or Z is an empty set, then $X\twoheadrightarrow Y|Z$ is called a trivial MVD; otherwise a non-trivial MVD.

The MVDs provide a necessary and sufficient condition for a relation to be decomposable into two of its projections without loss of information. The original relation is obtained as the natural join of the two projections. For example, $X\twoheadrightarrow Y$ holds for $R(X,Y,Z)$ iff every instance R is the natural join of projections $R[XY]$ and $R[XZ]$.

A relation schema is called a projection schema on X of $R(U)$, denoted by $R_k(X)$, if its instance is equal to $R[X]$. Let $R_k(X,Y,Z)$ be a projection schema of $R(W,X,Y,Z)$, where $W(\neq\emptyset)$ and $X\cup Y\cup Z(\neq\emptyset)$ are disjoint. The MVD $X\twoheadrightarrow Y|Z$ of $R_k(X,Y,Z)$ is called an embedded multivalued dependency (for short, EMVD)[2] of $R(W,X,Y,Z)$. If Y or Z is an empty set, then $X\twoheadrightarrow Y|Z$ is called a trivial EMVD; otherwise a non-trivial EMVD.

If an MVD $X\twoheadrightarrow Y$ holds for $R(U)$ and every $Y_R(x)$ contains at most one member, then the functional dependency (for short, FD) $X\rightarrow Y$ holds for $R(U)$.

It is important to distinguish a dependency satisfied in some specific instances from the one satisfied in a relation schema. If for some specific instance R of R(U) ($U \supseteq X \cup Y \cup Z$), $Z_R(x)=Z_R(xy)$ holds for any XY-value xy in R[XY], then we say $X \twoheadrightarrow Y|Z$ is valid in R.

Let D_i be a set of FDs, MVDs and EMVDs which hold for R(U). Let d be a single dependency. The dependency d is said to be implied by D_i iff d is valid in any instance of R(U) for which D_i holds. Two sets D_i and D_j of dependencies are said to be mutually derivable iff each element in D_i is implied by D_j and each element in D_j is implied by D_i, denoted by $D_i \sim D_j$. If D_i and D_j are not mutually derivable, then we denote it by $D_i \not\sim D_j$. If $D_i \sim D_j$, then D_j (D_i) is called a covering of D_i (D_j).

The binary relation $\sim$ on a family of sets of dependencies is an equivalence relation, that is, (1) $D_i \sim D_i$, (2) if $D_i \sim D_j$, then $D_j \sim D_i$, and (3) if $D_i \sim D_j$ and $D_j \sim D_k$, then $D_i \sim D_k$.

3. Properties of Embedded Multivalued Dependencies

In this section, we provide basic properties of EMVDs which are useful to handle the representability problem.

Fagin[2] and independently, Zaniolo[6] provided the following theorem in order to obtain a set of all the EMVDs implied by an MVD.

Theorem 1: [2,6] If an MVD $X \twoheadrightarrow Y|Z$ holds for R(X,Y,Z), then the EMVD $X \twoheadrightarrow Y'|Z'$ holds for R(X,Y,Z), where X, Y and Z are disjoint sets of attributes and $Y' \subseteq Y$, $Z' \subseteq Z$ and either Y' or Z' is a proper subset.

We note that the following results cannot be obtained from Beeri's complete set of inference rules for MVDs[7] since their inference rules are applicable only to a fixed, given relation schema.

Theorem 2: [8,9] Both the MVD $XY \twoheadrightarrow Z$ and the EMVD $X \twoheadrightarrow Y|Z$ hold for R(U) iff the MVD $X \twoheadrightarrow Z$ holds for R(U), where U is a set of attributes such that $U \supsetneq X \cup Y \cup Z$.

Corollary 1: [8,9] Let U, U', X and Y be sets of attributes such that $U \supset U' \supseteq X \cup Y$. Let $R_1(U')$ be a projection schema of R(U). Any MVD $X \twoheadrightarrow Y$ of $R_1(U')$ also holds for R(U) iff the MVD $U'-Y \twoheadrightarrow Y$ holds for R(U).

Theorem 3: [8,9] Assume that both $X_1 \twoheadrightarrow Y_1|Z_1$ and $X_2 \twoheadrightarrow Y_2|Z_2$ hold for R(U), where $U \supseteq X_1 \cup Y_1 \cup Z_1$, $X_1 \cup Y_1 = X_2 \cup Y_2 \cup Z_2$. X_1, Y_1 and Z_1 are disjoint, and X_2, Y_2 and Z_2 are also disjoint. Then, $X_2(Y_2-Y_1) \twoheadrightarrow Y_1 \cap Y_2|Z_1Z_2$ holds for R(U).

Example 1: Consider, for example, a bibliography relation schema which consists of attributes: paper-ID, author, affiliation, keyword and related-term. We assume that each author uniquely determines a set of the author's affiliations, each keyword uniquely determines a set of the keyword's related-terms, and each paper-ID uniquely determines a set of authors and a set of keywords. These constraints are represented as:

(1) $\{\text{author}\} \twoheadrightarrow \{\text{affiliation}\} | \{\text{paper-ID, keyword, related-term}\}$

(2) $\{\text{keyword}\} \twoheadrightarrow \{\text{related-term}\} | \{\text{paper-ID, author, affiliation}\}$

(3) $\{\text{paper-ID}\} \twoheadrightarrow \{\text{author}\} | \{\text{keyword}\}$.

By Theorem 1, Augmentation rule[7] and (2), we obtain the EMVD:

(4) $\{\text{paper-ID, keyword}\} \twoheadrightarrow \{\text{related-term}\} | \{\text{author}\}$.

By Theorem 2, (3) and (4), we obtain the EMVD:

(5) $\{\text{paper-ID}\} \twoheadrightarrow \{\text{keyword, related-term}\} | \{\text{author}\}$.

By Theorem 3, (1) and (5), we obtain the MVD:

(6) $\{\text{paper-ID}\} \twoheadrightarrow \{\text{keyword, related-term}\} | \{\text{author, affiliation}\}$.

In [10], Nicolas showed the first order logic formalization for FDs and MVDs. We show that it is also possible to express EMVDs as well-formed formulas (for short, wff) of logic. The first order logic formalization for EMVDs are useful from the following reasons: (a) As described in Section 4, EMVDs are important factors to consider the representability problem including MVDs and FDs. The first order logic will be a useful tool to investigate the properties of EMVDs, which are not completely known. (b) Non-dependency constraints also play an important role in the representability problem. As pointed out by Nicolas, the first order logic is useful to investigate the interactions between non-dependency constraints and dependency constraints, especially EMVDs.

For a relation schema R(U), $U=\{A_1,\ldots,A_n\}$, let X, Y, Z and W be disjoint subsets of U and $U' \cup W=U$, $X \cup Y \cup Z=U' \subsetneq U$. The EMVD $X \twoheadrightarrow Y|Z$ is valid in an instance R iff R[XYZ] is the natural join of its projections R[XY] and R[XZ].

Let us associate a variable x_i to each $A_i \in U$, a second variable x'_j to each $A_j \in U-X$, and a third variable x''_k to each $A_k \in W$. Let $U'=\{A_{i1},\ldots, A_{ip}\}$ $(0 \leq p < n)$, $Z \cup W=\{A_{j1},\ldots,A_{jq}\}$ $(0 \leq q \leq n)$, $Y=\{A_{j_{q+1}},\ldots,A_{j_r}\}$ $(0 \leq r-q \leq n)$, and $W=\{A_{k1},\ldots,A_{ks}\}$ $(0 \leq s \leq n)$.

Then, the wff WE corresponding to the EMVD $X \twoheadrightarrow Y|Z$ is as follows:

$$\text{WE:}\quad \forall x_{i1}\cdots\forall x_{ip}((\exists x'_{j1}\cdots\exists x'_{jq}\ R(u_1,\ldots,u_n) \land \exists x'_{j_{q+1}}\cdots\exists x'_{j_r}\exists x_{k1}\cdots\exists x_{ks} R(v_1,\ldots,v_n)) \xrightarrow[i]{} \exists x''_{k1}\cdots\exists x''_{ks} R(w_1,\ldots,w_n)).$$

Here, R is a name for an n-place predicate, and $\xrightarrow[i]{}$ denotes an implication. u_i is identical to x_i if $A_i \in X \cup Y$; otherwise x_i'. v_i is identical to x_i if $A_i \in U-Y$; otherwise x_i'. w_i is identical to x_i if $A_i \in U'$; otherwise x_i''.

Theorem 4: WE is true in an interpretation iff the instance R satisfies the EMVD.

According to the formalization above, for example, the EMVD $X_1 \rightarrow\rightarrow X_2 | X_3$ of $R(X_1,X_2,X_3,X_4)$ is expressed by the wff WE1 as follows:

$$\text{WE1: } \forall x_1 \forall x_2 \forall x_3 ((\exists x_3' \exists x_4' R(x_1,x_2,x_3',x_4') \wedge \exists x_2' \exists x_4 R(x_1,x_2',x_3,x_4)) \xrightarrow[i]{} \exists x_4'' R(x_1,x_2,x_3,x_4'')).$$

Both Theorem 1 and Theorem 2 can be proved by the first order logic formalization for MVDs and EMVDs, and automatic theorem proving techniques[11].

4. Update-Independent Representability of Decompositional Schema Design with MVDs

In this section, we provide a useful criterion for the representability for the decompositional schema design with MVDs and FDs. Properties of EMVDs obtained in Section 3 are used to provide conditions for a set of relation schemata to satisfy this criterion.

Fig. 1 shows several design methods and their criteria for the representability problem. Rissanen's result is summarized by two properties: lossless join property[4] and covering property for FDs. They are formulated by $R = \mathop{*}_{i=1}^{n} R[U_i]$ and $F \sim \bigcup_{i=1}^{n} F_i$, respectively, where $\{R_i(U_i);\ i \in \{1,\ldots,n\}\}$ is a set of relation schemata obtained from the universal (initial) relation schema $R(U)$, and each F_i is a covering of all the FDs that are defined on $R_i(U_i)$ for a given set F of FDs of R(U).

The lossless join property guarantees that both an initial schema and designed schemata have the same data, that is, the data of the initial schema is obtained by join operations from the data corresponding to the designed schemata. This property is automatically satisfied since Rissanen's design method is based on the decomposition approach[4].

The covering property for FDs guarantees that there exists no inter-relational FDs in the designed set of relation schemata. On the other hand, in Fagin's decomposition approach including MVDs, the criterion for having the same dependencies has not been studied.

Since Rissanen's result is based on only FDs, they cannot be directly

Fig.1 Several criteria for the representability problem

Methods	Synthetic (Bernstein)	Decomposition (Fagin)	Decomposition (Rissanen)	Decomposition (Authors)
Inputs	FDs	FDs+MVDs	FDs	FDs+MVDs (+EMVDs)
Assumption	uniqueness of FDs	universal relation	universal relation	universal relation
Same data	-	lossless join	lossless join	lossless join
Same dependencies	covering property for FDs	-	covering property for FDs	covering property for MVDs & FDs
Preserveness under updates	FD-preserved	-	FD-preserved	MVD(FD)-preserved

generalized for handling the representability problem with MVDs and FDs because of the following problems:

(A) Any FD used in the decomposition process always holds for the initial schema. On the other hand, when MVDs are also involved, we often use an EMVD $X \twoheadrightarrow Y|Z$ in a decomposition, such that neither MVD $X \twoheadrightarrow Y$ nor MVD $X \twoheadrightarrow Z$ holds for the initial relation schema. It requires to handle not only MVDs, but also EMVDs when MVDs are involved.

(B) Any FD $X \longrightarrow Y$, that is used to decompose $R(X,Y,Z)$ into $R_1(X,Y)$ and $R_2(X,Z)$, is embodied by the obtained schema $R_1(X,Y)$. On the other hand, any MVD $X \twoheadrightarrow Y$ ($X \not\rightarrow Y$, $Z \neq \emptyset$), that is used to decompose $R(X,Y,Z)$ into $R_1(X,Y)$ and $R_2(X,Z)$, is not embodied by $R_1(X,Y)$ since MVD $X \twoheadrightarrow Y|Z$ cannot be defined on $R_1(X,Y)$. It requires to handle not only dependencies embodied by relation schemata, but also those used in the decomposition.

(C) When only FDs are involved, any FD implied by a given set of FDs is preserved to be valid in the instance R whenever some R_i is updated, if the covering property is satisfied. On the other hand, some MVD is not preserved to be valid in R even if the covering property for MVDs and FDs (MVD-covering property) is satisfied.

Example 2: As for (A) and (B), consider, for example, the bibliography relation schema previously shown in Example 1. The initial relation schema

R(paper-ID,author,affiliation,keyword,related-term) can be decomposed into $\{R_1$(author,affiliation), R_2(keyword,related-term), R_3(paper-ID, author), R_4(paper-ID,keyword)$\}$ by using (1), (2) and (3).

The EMVD $\{$paper-ID$\} \twoheadrightarrow \{$author$\}|\{$keyword$\}$ is used in the decomposition although neither the MVD $\{$paper-ID$\} \twoheadrightarrow \{$author$\}$ nor the MVD $\{$paper-ID$\}$ $\twoheadrightarrow \{$keyword$\}$ holds for the initial relation schema. Furthermore, R_3(paper-ID,author) embodies only trivial EMVDs such as $\{$paper-ID$\} \twoheadrightarrow$ $\{$author$\} | \emptyset$.

Example 3: As for (C), consider, for example, another bibliography relation schema which consists of attributes: paper-ID, keyword, related-term and category such that

(1) $\{$keyword$\} \twoheadrightarrow \{$paper-ID$\}|\{$related-term,category$\}$

(2) $\{$category$\} \twoheadrightarrow \{$related-term$\}|\{$keyword,paper-ID$\}$.

We can prove that $\{(1),(2)\} \sim \{(3),(4),(5)\}$ such that

(3) $\{$related-term,keyword$\} \twoheadrightarrow \{$paper-ID$\}|\{$category$\}$

(4) $\{$keyword$\} \twoheadrightarrow \{$paper-ID$\}|\{$related-term$\}$

(5) $\{$category$\} \twoheadrightarrow \{$keyword$\}|\{$related-term$\}$.

By (3), (4) and (5), we can obtain $S_1 = \{R_1$(keyword,paper-ID), R_2(keyword, related-term), R_3(category,keyword), R_4(category,related-term)$\}$. We can easily verify that some deletion of tuples from R_2 would violate (5) in $(R_1 * R_2) * (R_3 * R_4)$.

Definition 1: Given a set of an initial relation schema $S_0 = \{R(U)\}$, let $S_k = \{R_1(U_1), \ldots, R_n(U_n)\}$ (k correspods to the subscript of D_k defined below) be a set of relation schemata obtained from R(U) by a decomposition. I_0 denotes a set of given FDs, MVDs and EMVDs that hold for R(U). D_k denotes a set of MVDs , EMVDs and FDs of R(U) that are actually used to obtain S_k in the decomposition. For each $i \in \{1, \ldots, n\}$, I_i denotes a covering of the FDs, non-trivial MVDs and non-trivial EMVDs defined on $R_i(U_i)$, that are implied by I_0.

Assumption 1: We assume that neither a trivial MVD nor a trivial EMVD is used in any decomposition. That is, any D_k does not contain such MVDs and EMVDs.

Assumption 2: We assume that for any S_k, each instance R_i of $R_i(U_i)$ in S_k is allowed to be updated according to only I_i.

Under Assumptions 1 and 2, our main problem for the representability is whether or not the given set I_0 of dependencies is at any time valid in R, which is a reconstructed instance of R(U) by join operations in the reverse order of the decomposition of R(U).

We define the criterion for the representability as follows:

Definition 2: S_k is said to have the Update-Independent(UI)-representability for S_0 iff any dependency implied by I_0 is valid in any instance of R(U) that is reconstructed in the reverse order of the decomposition. Here, S_k obeys both Assumptions 1 and 2.

Definition 3: S_k is said to satisfy the MVD-covering property iff $I_0 \sim (\bigcup_{i=1}^{n} I_i \cup D_k)$ holds. S_k is said to satisfy the MVD-preserving property iff any dependency belonging to $(\bigcup_{i=1}^{n} I_i \cup D_k)$ is valid in any instance of R(U) that is reconstructed in the reverse order of the decomposition. Here, S_k is assumed to obey Assumptions 1 and 2.

From Definitions 2 and 3, we immediately obtain the following theorem:

Theorem 5: S_k has the UI-representability for S_0 iff S_k satisfies both the MVD-covering property and the MVD-preserving property.

The MVD-preserving property is based on the following theorems:

Theorem 6: Let V,W,X,Y,Z be arbitrary disjoint sets of attributes. Assume that there are two relation schema $R_1(W,X,Y,Z)$, for which $WX \twoheadrightarrow Y|Z$ holds, and $R_2(V,X,Y,Z)$. $WX \twoheadrightarrow Y|Z$ is valid in R_1*R_2 for any instances R_1 and R_2 iff when $WX \not\rightarrow Y$ and $WX \not\rightarrow Z$, $X \twoheadrightarrow Y|Z$ holds for $R_2(V,X,Y,Z)$ (and consequently for $R_1(W,X,Y,Z)$). Here, we assume that each relation schema obeys only dependency constraints.

Note that in Theorem 6, each relation schema is assumed to obey only dependency constraints. If we allow non-dependency constraints, the necessary condition does not hold in Theorem 6. For example, let $R_1(W,X,Y,Z)$ obey the non-dependency constraint:
$|Y_{R_1}(w_1x_1)|=1$ or $|Z_{R_1}(w_1x_1)|=1$ holds for all WX-value w_1x_1 in any R_1. In this case, $WX \twoheadrightarrow Y|Z$ is always valid in any R_1*R_2 even if $WX \not\rightarrow Y$, $WX \not\rightarrow Z$ and $X \not\twoheadrightarrow Y|Z$.

Theorem 7: Let $V,W,X_1,X_2,Y_1,Y_2,Z_1,Z_2$ be arbitrary disjoint sets of attributes. Assume that there are two relation schemata $R_1(W,X_1,X_2,Y_1,Y_2,Z_1,Z_2)$, for which $X_1X_2 \twoheadrightarrow Y_1Y_2|Z_1Z_2$ holds, and $R_2(V,X_1,Y_1,Z_1)$. $X_1X_2 \twoheadrightarrow Y_1Y_2|Z_1Z_2$ is valid in R_1*R_2 for any instances R_1 and R_2 iff $X_1 \twoheadrightarrow Y_1|Z_1$ holds for $R_2(V,X_1,Y_1,Z_1)$ (and consequently for $R_1(W,X_1,X_2,Y_1,Y_2,Z_1,Z_2)$ when $X_1X_2 \not\rightarrow Y_1$ and $X_1X_2 \not\rightarrow Z_1$.

Theorem 8: Let $U(\neq\emptyset),V,W,X_1,X_2,Y_1,Y_2,Z_1,Z_2$ be arbitrary disjoint sets of attributes. Assume that there are two relation schemata $R_1(U,W,X_1,X_2,Y_1,Y_2,Z_1,Z_2)$, for which $X_1X_2 \twoheadrightarrow Y_1Y_2|Z_1Z_2$ holds, and $R_2(U,V,X_1,Y_1,Z_1)$. Then, $X_1X_2 \twoheadrightarrow Y_1Y_2|Z_1Z_2$ is valid in R_1*R_2 for any instances R_1 and R_2 if

[1] $X_1X_2 \rightarrow Y_1Y_2$ or $X_1X_2 \rightarrow Z_1Z_2$ holds or

[2] $X_1X_2 \rightarrow\!\!\rightarrow Y_1Y_2|Z_1Z_2U$ and $X_1 \rightarrow\!\!\rightarrow Y_1|Z_1U$ hold or

[3] $X_1X_2 \rightarrow\!\!\rightarrow Y_1Y_2U|Z_1Z_2$ and $X_1 \rightarrow\!\!\rightarrow Y_1U|Z_1$ hold.

From Theorems 7 and 8, the MVD-preserving property can be examined (partly) by the following basic algorithm:

Step 1: Let $R_i(U_i)$ and $R_j(U_j)$ $(i \neq j)$ belong to S_k such that $R_i(U_i)$ and $R_j(U_j)$ are obtained by decomposing a relation schema $R_m(U_iU_j)$. For each such $R_i(U_i)$ and $R_j(U_j)$, examine whether or not each dependency in I_i (I_j) is valid in $R_i * R_j$ by Theorems 7 and 8.

Step 2: If some dependency is proved not to be valid by Theorem 7, then S_k does not satisfy MVD-preserving property. If Theorem 8 is applied and the condition is not satisfied, then it is not known whether or not S_k satisfies the MVD-preserving property by this algorithm.

If any dependency in I_i and I_j is proved to be valid in $R_1 * R_2$, then

$$S_k = (S_k - \{R_i(U_i), R_j(U_j)\}) \cup \{R_m(U_iU_j)\}, \text{ and}$$

$$I_m = I_i \cup I_j \cup \{U_i \cap U_j \rightarrow\!\!\rightarrow U_i - U_j | U_j - U_i\}.$$

Step 3: Repeat Steps 1 and 2 until we obtain an initial relation schema.

For a given S_0 and I_0, there may exist D_i and D_j $(i \neq j)$ such that S_i does not satisfy the MVD-preserving property, but S_j does although both S_i and S_j satisfy the MVD-covering property. Fig.2 shows such an example, in which S_0 and I_0 are the same as Example 3.

Fig.2 Another decomposition of R(paper-ID,keyword,related-term, category).

$S_0 = \{$R(paper-ID,keyword,related-term,category)$\}$ $I_0 = \{(1),(2)\}$ in Example 3

R_1(keyword,paper-ID)　　R'(keyword,related-term,category)

R_2(category,related-term)　R_3(category,keyword)

$S_2 = \{R_1$(keyword,paper-ID), R_2(category,related-term), R_3(category,keyword)$\}$

$D_2 = \{\{\text{keyword}\} \rightarrow\!\!\rightarrow \{\text{paper-ID}\}|\{\text{related-term,category}\},$
$\{\text{category}\} \rightarrow\!\!\rightarrow \{\text{related-term}\}|\{\text{keyword}\}\}$

In this case, $I_0 \sim D_2$ and MVD-preserving property is also satisfied by S_2.

Acknowledgements: The authors are grateful to the colleagues in Yajima Laboratory, especially Mr. Y. Matsumoto for useful discussions.

This work is partly supported by the Science Foundation Grant of the Ministry of Education, Science and Culture of Japan.

References:

[1] Codd,E.F., "Further Normalization of the Data Base Relational Model", Courant Comput. Sci. Symposia, pp.34-64, May 1971

[2] Fagin,R., "Multivalued Dependencies and a New Normal Form for Relational Databases", ACM TODS, vol.2 , no.3, pp.262-278, Sept. 1977

[3] Beeri,C., Bernstein,P.A. & Goodman,N., "A Sophisticate's Introduction to Database Normalization Theory", Proc. of 4th International Conference on VLDB, pp.113-124, Sept. 1978

[4] Rissanen,J., "Independenct Components of Relations", ACM TODS, vol.2, no.4, pp.317-325, Dec. 1977

[5] Aho,A.V., Beeri,C. & Ullman,J.D., "The Theory of Joins in Relational Data Bases (Extended Abstract)", Proc. of 18th Annual Symposium on Foundations of Computer Science, pp.107-113, Oct. 1977

[6] Zaniolo,C., "Analysis and Design of Relational Schemata for Database Systems", Univ. of Calif., Computer Methodology Group Rep., UCLA-ENG-7669, July 1976

[7] Beeri,C, Fagin,R. & Howard,J.H., "A Complete Axiomatization for Functional and Multivalued Dependencies in Database Relations", Proc. of ACM-SIGMOD International Conference, pp.47-61, Aug. 1977

[8] Kambayashi,Y., Tanaka,K. & Yajima,S., "Storage Anomalies and Normalization Problems in Relational Databases" (in Japanese), IECEJ SIGAL-Record AL77-71, Jan. 1978

[9] Tanaka,K., Kambayashi,Y. & Yajima,S., "Properties of Embedded Multivalued Dependencies in Relational Databases", Kyoto University, Yajima Lab. Research Report ER78-03, Dec. 1978

[10] Nicolas,J.M., "First Order Logic Formalization for Functional, Multivalued and Mutual Dependencies", Proc. of ACM-SIGMOD International Conference, pp.40-46, May 1978

[11] Chang,C.L. & Lee,R.C.T., "Symbolic Logic and Mechanical Theorem Proving", Computer Science and applied mathematics, Academic Press, Inc., 1973